Copyright © 1998, American Society for Microbiology. All Rights Reserved.

Multiple Orientation-Dependent, Synergistically Interacting, Similar Domains in the Ribosomal DNA Replication Origin of the Fission Yeast, *Schizosaccharomyces pombe*

SOO-MI KIM AND JOEL A. HUBERMAN*

Department of Genetics, Roswell Park Cancer Institute, Buffalo, New York 14263

Received 3 June 1998/Returned for modification 3 August 1998/Accepted 19 August 1998

Previous investigations have shown that the fission yeast, Schizosaccharomyces pombe, has DNA replication origins (500 to 1500 bp) that are larger than those in the budding yeast, Saccharomyces cerevisiae (100 to 150 bp). Deletion and linker substitution analyses of two fission yeast origins revealed that they contain multiple important regions with AT-rich asymmetric (abundant A residues in one strand and T residues in the complementary strand) sequence motifs. In this work we present the characterization of a third fission yeast replication origin, ars3001, which is relatively small (~570 bp) and responsible for replication of ribosomal DNA. Like previously studied fission yeast origins, ars3001 contains multiple important regions. The three most important of these regions resemble each other in several ways: each region is essential for origin function and is at least partially orientation dependent, each region contains similar clusters of A+T-rich asymmetric sequences, and the regions can partially substitute for each other. These observations suggest that ars3001 function requires synergistic interactions between domains binding similar proteins. It is likely that this requirement extends to other fission yeast origins, explaining why such origins are larger than those of budding yeast.

To understand the molecular mechanisms of initiation and regulation of eukaryotic DNA replication, it is helpful to study the *cis*-acting sequences (replicators) essential for origin function. Identification and characterization of replicators in the budding yeast, *Saccharomyces cerevisiae*, were facilitated by the discovery that DNA sequences serving as replicators in chromosomes also serve as replicators in plasmids (reviewed in reference 30). These origin sequences are called autonomously replicating sequence (ARS) elements, because they permit plasmids to replicate autonomously in yeast cells.

Plasmids bearing ARS elements and selectable markers can transform S. cerevisiae cells at a high frequency (15, 38), permitting a plasmid transformation assay that has been used to characterize the sequence requirements for origin function. ARS elements have two essential components: a short A domain of about 20 bp, which contains a \geq 9-of-11 match to the 11-bp ARS consensus sequence (ACS) (4, 30), and a broad (\sim 100-bp) flanking region, called the B domain, 3' to the consensus T-rich strand. The B domain consists of two or three additional important sequence motifs (16, 23, 31, 39). One of the important sequences in the B domain, called B1, cooperates with the A domain to form a binding site for the origin recognition complex (ORC), the putative initiator protein (2, 32, 33). Other possible functions for the B domain include serving as a DNA unwinding element (28, 41), enhancing origin activity by serving as a binding site for transcription factors (23, 42), and interacting with a single-stranded DNA binding protein (24).

Replication origins in most other eukaryotes are poorly understood, mainly due to lack of unambiguous techniques for characterizing them. However, in the fission yeast, *Schizosac*-

charomyces pombe, which is evolutionarily distant from *S. cerevisiae* (3), chromosomal DNA sequences with properties similar to ARS elements of budding yeast have been identified (25, 26, 34, 40, 44), and some of them have been shown to correspond to chromosomal replication origins (8, 35, 43). Because *S. pombe* is in some respects more similar to other eukaryotic organisms than is *S. cerevisiae* (reviewed in reference 45), it is possible that further study of ARS elements in *S. pombe* will provide information useful in understanding replication origins in many other eukaryotic organisms.

S. pombe ARS elements are AT rich, like those of S. cerevisiae, but are generally bigger (500 to 1500 bp). Two S. pombe ARS elements, ars1 (5) and ars3002 (7), have been studied in some detail. Deletion and linker substitution analyses indicate that these ARS elements contain at least one (for ars1) or two (for ars3002) essential modules and some additional important modules, and each of the essential modules contains critical sequence elements extending for 20 to 30 bp. The critical sequences are all AT rich and asymmetric, in the sense that A residues are clustered on one strand while T residues are clustered on the complementary strand.

To test whether these features are common to other *S. pombe* ARS elements, we chose to study the ribosomal DNA (rDNA) ARS element, which we have previously mapped to the nontranscribed spacer in the rDNA repeats (35). Because there are 100 to 150 copies of the rDNA repeat in the *S. pombe* genome, the rDNA ARS element is by far the most abundant ARS element in the genome.

We have previously shown that a 2.3-kbp *BamHI-KpnI* restriction fragment within the rDNA repeat (see Fig. 1) exhibits as much ARS activity as a restriction fragment containing the entire rDNA repeat and that the 2.3-kbp fragment contains all detectable rDNA initiation sites (see the gray box in Fig. 1 and reference 35). The ARS element within this 2.3-kbp fragment was designated *ars3001* according to the four-digit *S. pombe* ARS-naming convention (8), because it was the first ARS element to be discovered in chromosome III (9, 40).

^{*} Corresponding author. Mailing address: Department of Genetics, Roswell Park Cancer Institute, Elm & Carlton Streets, Buffalo, NY 14263-0001. Phone: (716) 845-3047. Fax: (716) 845-8126. E-mail: huberman@acsu.buffalo.edu.

In this report, we describe the results of systematic mutagenesis of *ars3001*. These results identify three domains that are essential for function. Each of these domains contains important sequences which share similarities with those detected in the two earlier studies but are not equivalent to the ACS of *S. cerevisiae* ARS elements. Domain substitution experiments indicate that the three domains are largely orientation dependent and can partially substitute for each other.

MATERIALS AND METHODS

Strains and media. Escherichia coli DH5 α cells (Life Technologies) were used for cloning plasmids containing mutated ars3001 DNA. The S. pombe strain ura4-D18 (ura4-D18 leu1-32 end1 h⁻) (12) was used as the recipient strain for transformation assays. Cells were grown in EMM (27) supplemented with 150 mg (each) of uracil and leucine per liter when not under selection or 150 mg of leucine per liter when under selection for uracil prototrophy.

Generation of progressive deletions and determination of nucleotide sequence of the 2.3-kbp BamHI-KpnI restriction fragment. Plasmid pRS306 (or 406): rDNA-2.3k was constructed by ligating the 2.3-kbp BamHI-blunt-ended-KpnI fragment (see Fig. 1) between the *Bam*HI and *Sma*I sites in the multiple cloning site of the vector pRS306 or pRS406 (37). Unidirectional sequential deletions from both ends of the insert were generated with exonuclease III as described previously (13) with modifications suggested by Stratagene (protocol for the pBluescriptII exonuclease III-mung bean DNA sequencing system; Stratagene). For deletions starting at the BamHI site, plasmid pRS406:rDNA-2.3k was digested with BamHI, to generate recessed 3' ends susceptible to exonuclease III attack, and also with SacI, to generate exonuclease III-resistant protruding 3' ends. For deletions in the opposite orientation, plasmid pRS306:rDNA-2.3k was digested with EcoRI to generate susceptible ends. To generate resistant ends, treatment with HindIII was followed by treatment with Klenow polymerase and $\alpha\text{-thio-deoxynucleoside}$ triphosphates. Subsequent incubation with exonuclease III for increasing times resulted in progressive resection of the insert while leaving the SacI or α-thio-HindIII end intact. Mung bean nuclease was then used to generate flush ends. The plasmids were recircularized by blunt-end ligation and then used to transform DH5α cells. Minipreps of individual plasmid clones from each time point were used to select clones with deletions of increasing sizes in increments of 100 to 250 bp. These clones were then used, in combination with vector-specific primers, to determine the nucleotide sequence of both strands of the 2.3-kbp BamHI-KpnI fragment.

Plasmid construction. To obtain the plasmid pura4script:rDNA-573, the exonuclease III deletion construct, K4 (see Fig. 2A), was cut with FokI and treated with Klenow polymerase. The resulting rDNA-containing 1.86-kbp fragment was gel purified and cut with ClaI to generate a 583-bp fragment (the desired 573-bp fragment plus 10 extra bp of vector sequence). The gel-purified 583-bp fragment was ligated between the ClaI and SmaI sites in the multiple cloning site of the vector pura4script (8), which contains the S. pombe ura4 gene, thus permitting selection in the ura4-D18 strain.

Generation of \sim 60-bp internal deletions. All the \sim 60-bp deletion constructs except $\Delta 3$ (see Fig. 3A) were made according to the protocol for the MORPH Site-Specific Plasmid DNA Mutagenesis Kit (5 Prime-3 Prime, Inc.). The primers have 13- to 18-bp flanking sequences at both sides of the region to be deleted and replaced by the EcoRI-Bg/II linker, GGAATTCCGAAGATCTTC, at the position of the deletion. They were annealed to the template plasmid, pura4script: rDNA-573, prepared from *E. coli* (methylated). Subsequent incubation with T4 DNA polymerase and T4 DNA ligase resulted in a mixture of nonmutagenized template (both strands methylated) and mutagenized (methylated template strand and nonmethylated replacement strand) plasmids. The nonmutagenized template plasmids were then fragmented by DpnI, and the mutagenized plasmids were transformed into an E. coli mutS strain, in which the methylation-specific repair system is inactive. The desired constructs were identified by susceptibility to digestion with EcoRI or BglII. For $\Delta 3$, the two-step PCR method described by Dubey et al. (7) was used (see below). The external primers, rightward-pointing (forward) and leftward-pointing (reverse), correspond to sequences within the multiple cloning site outside the 573-bp insert. The internal (forward and reverse) primers flanking the region to be deleted (region 3) had the EcoRI-BglII linker and the EcoRI linker, respectively, at their 5' ends. All primer sequences are available upon request.

Generation of 10-bp linker substitutions. The procedure employed was similar to that of Dubey et al. (7). All internal (forward and reverse) primers were designed with stretches of 17 to 24 nucleotides (depending on requirements for specific PCR) corresponding to the flanking sequences of the regions to be substituted (indicated by horizontal lines in Fig. 5). In addition, the primers had linkers containing AvaI restriction sites (CCCCGGGGGG) at their 5' ends. External primers were based on sequences in the vector multiple cloning site and were designed so that the XbaI and XhoI sites within the multiple cloning site would be included in the final PCR products. The forward external primer was the same as that described above (previous paragraph), and the reverse external primer was designed to delete the AvaI site located within the multiple cloning site. To generate linker substitutions, internal primers flanking regions to be

substituted were amplified in combination with appropriate external primers. The resulting PCR products were digested with AvaI and then ligated together. Then, the ligation products were further amplified with the paired external primers, digested with XbaI and XhoI, and cloned between the XbaI and XhoI sites of pura4script. The consequence of these manipulations was replacement of 10 bp of S. pombe sequence by the 10-bp AvaI linker. All primer sequences are available on request.

Generation of domain substitutions and inversions. To facilitate constructing substitutions or inversions, each of the three domains was first deleted and replaced by the EcoRI-Bg/II linker (see Fig. 6), according to the same two-step PCR and cloning strategy used for the constructions of $\Delta 3$ and the 10-bp linker substitutions. Then, the desired domain (as a PCR product flanked by an EcoRI site at one side and a Bg/II site on the other side, depending on the intended orientation) was simply ligated between the EcoRI and Bg/II sites in the appropriate deletion mutant. Thus, each mutant contains the desired substitution or inversion flanked by two linkers. All primer sequences are available on request.

Transformation of *S. pombe* cells. *S. pombe* D18 (12) cells lacking the *ura4* gene were transformed (10) with equal amounts of DNA from the plasmids under test and then grown under selection for uracil prototrophy. After 5 to 6 days, plates were scanned, and the number and mean size of colonies were calculated by using an image-processing program that permitted objective discrimination between the larger colonies and the smaller background colonies produced by vector alone.

Analysis of DNA sequences. To calculate free energies of unwinding, we employed the Thermodyn program (28) with a sliding 100-bp window. MacVector software (Oxford Molecular Group) was employed to search for the following consensus sequences (in which W stands for A/T, R stands for A/G, and Y stands for T/C): M, WRTTTATTTAW (1 mismatch allowed) (25); Z, WWTTWTTTWTTT (1 mismatch allowed) (46); C, TTGTATTTTAATTTGTATTTTTTGT AATTT (10 mismatches allowed) (5); and D, WTWTWTTTTTTTTTTTTTTTTT A (3 mismatches allowed). Consensus sequence D has not previously been published. It is based on the 20-bp critical sequence in Δ10 of ars3002 (7). Using MacVector software, we first searched the database consisting of all known S. pombe ARS element sequences for ≥12-of-20 matches to the ars3002 critical sequence. Then, we used the Consensus program of the Genetics Computer Group software package to develop a consensus from the matches that we found. The resulting consensus is shown above as sequence D.

Nucleotide sequence accession number. The sequence of the 2.3-kbp *Bam*HI-*Kpn*I fragment described in this work has been deposited in GenBank under accession no. AF040270.

RESULTS

Improved localization and characterization of ars 3001. Our previous studies (35) indicated that all S. pombe rDNA ARS activity is localized within a 2.3-kbp BamHI-KpnI fragment (Fig. 1). To better localize ars 3001 activity within this fragment, we employed progressive exonuclease III deletion from both ends of the fragment. The results of assays for ARS activity in selected deletions from each set are shown in Fig. 2A. The left and right boundaries for ars3001 defined by the data in Fig. 2A are sharp compared to those of most other S. pombe ARS elements studied so far (e.g., ars3002, ars3003, and the right boundary of ars1, all of which show a gradual decrease of ARS activity as the size of the deletion increases [5, 7, 46]). However, the left boundary of ars1 is similarly sharp (5). Even though the boundaries of ars3001 appear sharp, the tested deletions were all unidirectional, leaving open the possibility that a defect generated by a deletion in the left side of the ARS element could be compensated by a region preserved on the right side of the ARS element and vice versa. To test this possibility and to obtain a small fragment with full ARS activity for further analyses, we subcloned several fragments (Fig. 2B) encompassing the boundaries defined in Fig. 2A, and we tested these small fragments for ARS activity. As shown in Fig. 2B, all the tested clones showed ARS activity comparable to that of the full 2.3-kbp fragment. Since the smallest fragment (573 bp) includes the essential region defined in Fig. 2A (gray box), we used this fragment for further studies.

Compared to previously studied strong (ars2-2 and ars2-1 [43]) and weak (ars3002 and ars3003 [17]) S. pombe ARS elements, the 573-bp fragment showed moderately strong ARS activity (18), which is consistent with our observation that the chromosomal rDNA origin is moderately efficient (35). In ad-

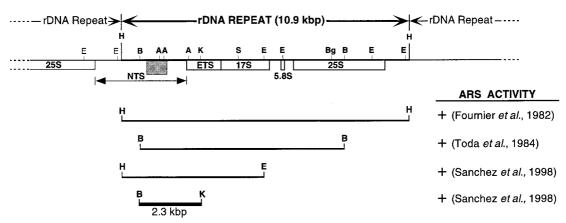


FIG. 1. Summary of previous studies localizing ARS activity in the rDNA repeat. (Top) The rDNA repeat (10.9 kbp) consists of the 3.46-kbp nontranscribed spacer (NTS) and the transcribed region containing the external transcribed spacer (ETS); the 17S, 5.8S, and 25S rRNA genes; and the internal transcribed spacers (ITS1 and ITS2, not labeled on the figure) which flank the 5.8S gene. Restriction enzyme sites are mapped based on the data from references 1, 19, 21 (GenBank accession no. Y09256) and 36 and this work (GenBank accession no. AF040270). The gray box shows the replication initiation zone where bubble arcs were detected by two-dimensional gel analysis (35). Abbreviations: A, AluI; B, BamHI; Bg, BgIII; E, EcoRI; H, HindIII; K, KpnI; S, SalI. (Bottom) Restriction fragments in the rDNA repeat previously shown to have ARS activity. The presence of ARS activity is indicated by a plus symbol. The 2.3-kbp BamHI-KpnI fragment (thick line) is the smallest tested fragment with full ARS activity.

dition, a plasmid containing the 573-bp fragment segregated stably as an unrearranged monomer for at least 20 cell generations (18).

Effects of ~60-bp internal deletions. The set of exonuclease III deletion clones that we had constructed permitted us to obtain the nucleotide sequence of the 2.3-kbp restriction fragment. Availability of the nucleotide sequence allowed us to conduct a higher-resolution mutational analysis of the ARS-containing 573-bp fragment.

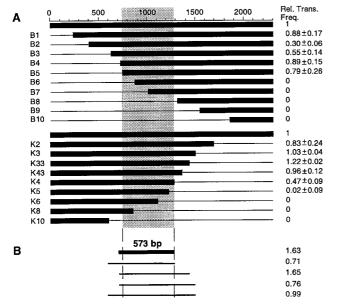


FIG. 2. Localization of ARS activity within the 2.3-kbp fragment. (A) The map shows sets of nested deletions either from the BamHI site (B1 to B10) or from the KpnI site (K2 to K10). Thin lines represent the portion deleted from the 2.3-kbp fragment (black bar at the top of the map). Transformation frequencies (Rel. Trans. Freq.) of the deletion constructs (means \pm standard deviations) are represented relative to that of the 2.3-kbp wild-type fragment. The gray box represents the minimal region showing significant ARS activity, as defined by external deletions. (B) The lines show the tested smaller fragments encompassing the minimal region defined in panel A. The thick line represents the smallest fragment (573 bp) showing a transformation frequency comparable to that of the complete 2.3-kbp fragment.

As the first step toward characterization of the sequence elements important for ARS activity within the 573-bp fragment, we introduced relatively large deletions (52 to 66 bp) (Fig. 3A) throughout the fragment and then tested the deletion constructs for transformation frequency. As shown in Fig. 3B, deletion of region 3 or 9 essentially eliminated ARS activity and deletion of region 2 or 6 significantly reduced it. Deletion of region 4, 5, or 7 moderately affected ARS activity.

Figure 3B presents measurements of both transformation frequency and average colony size. When cells are grown on selective medium, as they are during ARS activity assays, cell growth rate can be limited by plasmid replication rate. Plasmids bearing efficient ARS elements can replicate rapidly, and the cells containing them can produce large colonies. Consistent with these expectations, we found that the mean areas of colonies calculated by computer image analysis were roughly proportional to transformation frequency values (Fig. 3B). In a separate study on the ARS elements of the *ura4* origin region (17), we have shown that relative transformation frequencies and colony sizes of isolated ARS elements in plasmids correlate with origin function in the chromosome, as measured by the ratio of replication intermediates resulting from de novo initiation to those resulting from passive replication.

Since DNA unwinding can, in some cases, be a rate-limiting step in origin function (28, 41), we compared the results of our functional analysis (Fig. 3B) with the free energy of unwinding calculated for each position in the nucleotide sequence by using the Thermodyn program (28) employing a window size of 100 bp (thin black line in the ars3001 diagram in Fig. 4). Interestingly, the two regions with the lowest free energy of unwinding correspond to the two regions (3 and 9) which were most affected by the \sim 60-bp deletion (Fig. 3B; Fig. 4).

We also took advantage of the nucleotide sequence to search within the 573-bp fragment for sequence motifs previously detected by our laboratory (7, 46) and others (5, 25) as relatively abundant in *S. pombe* ARS elements. We shall refer to these as *S. pombe* ACSs, but it is important to realize that they are not equivalent to the ACS of *S. cerevisiae*. Prior to this investigation, it was known only that they are relatively abundant in *S. pombe* ARS elements and that two of them (5, 7) are found in particular sequences critical for ARS activity. Whether any of them has general significance for ARS func-

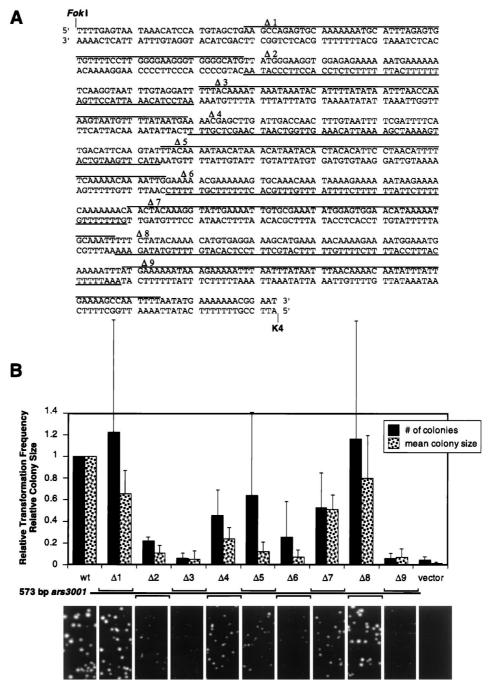
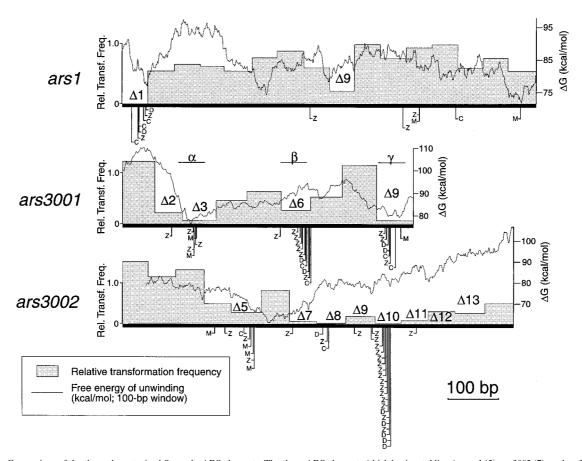


FIG. 3. Effects of \sim 60-bp internal deletions on the ARS activity of the 573-bp version of ars3001. (A) The nucleotide sequence of the 573-bp version of ars3001 is shown with the positions of each deletion region ($\Delta 1$ to $\Delta 9$). The left endpoint is a Fok1 cut site (see Materials and Methods) and the right endpoint is the K4 deletion point (Fig. 2). Over- and underlines are used to distinguish the boundaries of neighboring deletion regions. The size of each deletion is within the range 52 to 66 bp. The leftmost 28 nucleotides and the rightmost 19 nucleotides of the 573-bp fragment were not included in a deletion region. Instead, these sequences served as templates for primer design. The leftmost sequences are located beyond the left boundary defined by nested deletions (Fig. 2) and are thus not expected to be important for activity of the 573-bp fragment. The importance of the rightmost 19 nucleotides remains uncertain. (B) At the top are shown the transformation frequency and colony area relative to those of the 573-bp wild-type ars3001. Black bars and stippled bars represent values calculated for four independent experiments. Error bars show standard deviations. The vector alone (pura4script [7]) was transferred into S. pombe cells as a negative control, wt, wild type. In the middle is shown a map of the deletion regions, indicated by brackets. $\Delta 1$ to $\Delta 9$ correspond to deletion regions 1 to 9. At the bottom, pictured at the same magnification, are portions of petri plates from one of the experiments summarized in the chart at 6 days after transformation.

tion was unknown. It is striking, therefore, that all of the consensus sequence matches in *ars3001* (marked Z, D, M, and C in the *ars3001* diagram in Fig. 4) are localized in regions 2, 3, 6, and 9, the regions whose deletions most reduced ARS activity.

Effects of 10-bp linker substitutions. To obtain a higher-resolution picture of the nucleotide sequences important for ARS function in regions 2, 3, 6, and 9, we employed linker substitution with a 10-bp GC-rich linker (C_5G_5) . The linker was substituted for each of the overlined or underlined 10-bp nu-



cleotide sequences (Fig. 5), and the effect of the substitution on transformation was measured. Note that the results are displayed on a logarithmic scale.

In regions 2 and 3 (Fig. 5A), the linker substitutions at positions 2-6 and 3-3 decreased ARS activity about fourfold, and approximately twofold effects were created by the two intervening substitutions (3-1 and 3-2). However, none of the 10-bp substitutions in regions 2 and 3 inhibited ARS activity to the same extent as the \sim 60-bp deletion of region 2 or 3 (Fig. 3B). Thus, the stimulatory effects of regions 2 and 3 on ARS function may be due to redundant function of several short sequence stretches within these regions. This result is similar to but less extreme than that obtained in a study of ars1 (5), where a 50-bp deletion of the region at the extreme left end of ars1 decreased activity nearly 200-fold, yet none of the 10-bp linker substitutions in this region significantly affected ARS activity. Five of the six consensus sequence matches in regions 2 and 3 (lower portion of Fig. 5A) are located in the area most affected by linker substitution mutations (3-1 to 3-3), suggesting the importance of these consensus sequences for the stimulatory effect of region 3.

Surprisingly, linker substitution at position 6-4 reduced transformation frequency \sim 15-fold (Fig. 5B), even though deletion of the entire 54 bp of region 6 caused only an \sim 5-fold reduc-

tion in transformation frequency (Fig. 3B). The more extreme effect of the smaller linker substitution suggests that malfunction of region 6 may interfere with ARS activity to a greater extent than complete loss of region 6. In addition to linker substitution at position 6-4, that at 6-3 also has a significant effect (about threefold). The substitutions of 6-3 and 6-4 colocalize with several consensus sequences, lending further support to the hypothesis that these sequences have biological importance.

One linker substitution in region 9, that at 9-5, reduces ARS activity ~30-fold (Fig. 5C) and thus accounts for the effect (~30-fold) of deleting all 66 bp of region 9 (Fig. 3B). Two other linker substitutions (9-1 and 9-2) also have pronounced effects (about 10-fold) and colocalize with a cluster of consensus sequence motifs, adding to the evidence from regions 3 and 6 that clustered consensus sequences contribute to ARS function.

Effects of domain deletion, inversion, and substitution. The fact that clustered consensus sequences appear to be important for ARS activity in regions 3, 6, and 9 (Fig. 3B, 4, and 5) raises the question of what the function(s) of these sequences might be. The consensus sequence of *S. cerevisiae* ARS elements (the ACS) is an important part of the ORC binding site (2, 6). It is known that inversion of the *S. cerevisiae* ACS inactivates ARS activity (14). We wondered, therefore, what might be the ef-

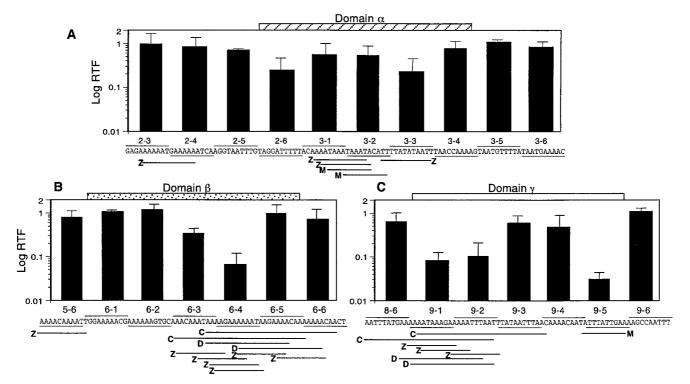


FIG. 5. Effects of 10-bp linker substitution mutations within deletion regions 2 to 3 (A), 6 (B), and 9 (C) on ars3001 activity. The nucleotides indicated by the horizontal lines were replaced by the 10-bp linker, CCCCCGGGGG, which contains the AvaI restriction site. Each number (2-3 to 9-6) marked above the sequence corresponds to the position of a linker substitution; e.g., 2-3 corresponds to the third linker substitution from the left end of deletion region 2, etc. Note that average transformation frequencies of four to six independent experiments relative to those of the 573-bp wild type are indicated on a logarithmic scale (Log RTF). Error bars show the standard deviations. At the bottom of each nucleotide sequence are shown the positions of matches to the Z (46), M (25), C (5), and D (see Materials and Methods) S. pombe ARS consensus motifs (see also Fig. 4). In each case, the distinguishing letter is at the 3' end of the T-rich strand of the consensus motif. The boundaries of 50-bp domains α , β , and γ are indicated by striped, dotted, and open boxes, respectively.

fects of inverting the important clustered-consensus-sequencecontaining regions in *ars3001*. We also wondered whether the contributions to ARS activity by regions 2, 3, 6, and 9 are unique or redundant, and we thought we might be able to distinguish between these possibilities by testing the effects of substituting the important sequences for each other.

To facilitate inversion and substitution experiments, we defined three domains covering the sequences identified as most important by linker substitution in regions 2, 3, 6, and 9 (Fig. 5). Since we do not know the role of the spacing between the domains, we set the domain sizes to be of equal length (50 bp), sufficient to include all of the 10-bp stretches identified as most important by linker substitution, even in region 9. We have arbitrarily called these domains α , β , and γ in order from left to right in *ars3001* (Fig. 4 to 7).

The first step in constructing the inversion and substitution mutations was the creation, by PCR mutagenesis, of precise deletions of each of the three domains, leaving a 20-bp linker pair (consisting of a 10-bp EcoRI linker plus a 10-bp BgII linker) in each of their places. To construct other mutations, we then simply inserted the desired domain, in the intended orientation, between the EcoRI and BgII sites in the appropriate deletion mutant. Thus, each mutant we constructed contains the desired substitution or inversion flanked by two linkers.

Since we had not previously tested the combined effects of two linkers harboring different restriction enzyme sites and these linkers were different from those used in the previous 10-bp linker substitution analysis (Fig. 5), we measured the effects of the new linkers in two ways. First, we confirmed that the linkers did not have ARS activities of their own. The constructs in which domains α , β , and γ were deleted and replaced by the 20-bp linker pair lacked significant ARS activity (Fig. 7 [$\mathbf{d}\alpha$, $\mathbf{d}\beta$, and $\mathbf{d}\gamma$]), as anticipated. Second, we tested the effects of substituting each domain by itself, thus generating constructs which differed from the wild-type ars3001 only in having two linkers flanking domain α , β , or γ (Fig. 6). As is evident from Fig. 6, the two linkers did not decrease transformation frequency, but they did decrease colony size, especially when they flanked domain β or γ . We used these results as controls for further inversion and substitution experiments.

Each domain was then replaced by other domains in normal or inverted orientation. All measurements were relative to the controls in Fig. 6 (data redisplayed in Fig. 7 as $\alpha s \alpha$, $\beta s \beta$, and $\gamma s \gamma$). The results (Fig. 7) show that orientation is critical for the operation of domains α and β (i α , α si β , α si γ , i β , β si α , β si γ) but less important for domain γ (i γ , γ si α , γ si β). In its normal orientation, domain α can substitute for domain β or γ $(\beta s\alpha, \gamma s\alpha)$, but neither domain β nor domain γ $(\alpha s\beta, \alpha s\gamma)$ can efficiently replace domain α . Domain γ can replace domain β $(\beta s \gamma)$, but domain β cannot efficiently replace domain γ $(\gamma s \beta)$. Thus, these observations suggest an order of domain functional capability for replacing other domains (from most to least capable: domain α , domain γ , domain β). The fact that some domains can substitute for others indicates that the functions of all three domains are, at least in some part, similar, in spite of any limitations imposed by our rather arbitrary definitions of the domains.

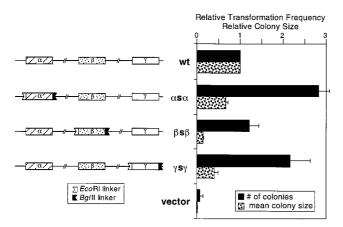


FIG. 6. Linker substitution at both ends of domains α , β , and γ reduces colony size but not transformation frequency. The combined effects on ARS activity of the two different linkers flanking each domain were measured as controls for further experiments (Fig. 7). On the left are diagrams of the constructs, α (in which domain α is replaced by α itself), β s β (β by β), γ s γ (γ by γ), and the 573-bp wild-type (wt) and vector controls. Average transformation frequencies and colony sizes from two independent experiments, relative to those of the wild type, are indicated (with experimental ranges [error bars]) on the right.

DISCUSSION

The experiments described in this paper provide the first high-resolution genetic analysis of *ars3001*, the ARS element in *S. pombe* rDNA. Because the ribosomal genes are repeated 100- to 150-fold in the haploid genome (22) and because initiation sites for rDNA replication colocalize with *ars3001* (35), it is likely that *ars3001* contains the *cis*-acting sequences for the most abundant replication origins in *S. pombe* chromosomes.

We found that full ars3001 activity is contained within a 573-bp stretch (Fig. 2), making ars3001 the shortest of the well-characterized S. pombe ARS elements (Fig. 4). Internal deletion scanning and linker substitution mutagenesis revealed three regions within ars3001 that are especially important for its activity (Fig. 3 to 5). Based on linker substitution mutagenesis within these three regions (Fig. 5), 50-bp domains α , β , and γ were defined to include the most important sequences in each region (Fig. 4 to 7). Each of the domains is essential for ARS activity (Fig. 7). Domains α and β (and to a lesser extent domain γ) are sensitive to orientation (Fig. 7). Domain α can substitute for domains β and γ , and domain γ can replace domain β (Fig. 7). Thus, neither domain β nor domain γ performs a unique role.

What is the significance of these observations for our understanding of replication origin function in *S. pombe?* In attempting to answer this question, it is important to consider what has been learned from earlier studies of *ars1* (5) and *ars3002* (7) as well as from the present study of *ars3001*. Results from all three investigations are summarized in Fig. 4. In the next several paragraphs we discuss the major conclusions that can be drawn from these investigations.

Internal redundancy in *S. pombe* ARS elements. In Fig. 4, the thick black bars show (all on the same scale) the minimal ARS elements as defined by external deletion studies. Above the black bars are light gray bar charts showing the results of deleting the indicated 50- to 60-bp segments of each ARS element. In each series, the individual deletions are numbered consecutively from left to right across the ARS element ($\Delta 1$, $\Delta 2$, etc.), but Fig. 4 displays labels only for those deletions having the most serious effects. Some deletions have very serious consequences for ARS activity, while others have negli-

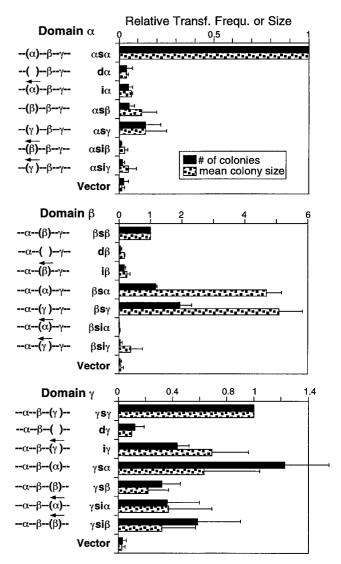


FIG. 7. Effects of deletions, substitutions, and inversions of the three domains on ARS activity. The transformation frequency (Transf. Frequ.) and colony size of each mutation were compared to those of the proper control construct (α s α for all domain α mutations, β s β for domain β , and γ s γ for domain γ [Fig. 6]). Average values of two independent experiments are shown with data ranges (error bars). Abbreviations: XsY, domain X replaced by domain Y; dX, deletion of X; iX, inversion of X; XsiY, domain X replaced by inverted orientation of domain Y.

gible impact. How should the different effects of these deletions be interpreted?

We and others (5, 7, 46) have pointed out that *S. pombe* ARS elements frequently contain redundant regions important for their activity. For example, although individual deletion of none of the rightmost three 50-bp segments of *ars1* reduces transformation frequency more than 2-fold, deletion of all three together reduces transformation frequency ~50-fold (5). Thus, the sequences important for ARS function within the rightmost 150 bp of *ars1* are redundant; only by deleting enough of them is their importance revealed. The lesson for interpretation of the 50- to 60-bp deletion results summarized in Fig. 4 is clear. Segments whose individual deletion seriously impairs ARS activity probably play unique roles in ARS function; segments whose deletions do not seriously inhibit ARS activity may also make important contributions; however, these

contributions may be redundant with contributions made by other segments.

The fact that ars3001 is the shortest of the well-characterized S.~pombe ARS elements suggests that it is less internally redundant. Perhaps that is why we could detect three important 50-bp domains, α , β , and γ (Fig. 4 to 7), each of which proved essential when replaced by a 20-bp linker pair (Fig. 7). This hypothetical reduced redundancy might also explain why 10-bp linker substitution mutations within domains α , β , and γ had measurable effects (Fig. 5) while linker substitutions as short as 10 bp had no significant effect even within the essential segment 1 of ars1 (5).

Because deletion scanning of ars1 detected only one essential segment ($\Delta 1$ [Fig. 4]), Clyne and Kelly (5) proposed that ars1 is organized similarly to S. cerevisiae ARS elements, with one domain containing the essential ACS and an essential flanking domain within which internal deletion and linker substitution mutations have reduced effects. The results subsequently obtained for ars3001 (this study) and ars3002 (7) suggest that such a close parallel between S. pombe and S. cerevisiae ARS element organization does not generally hold true. Both ars3001 and ars3002 contain multiple segments whose deletion seriously impairs ARS activity, and ars3002 contains two noncontiguous segments (8 and 10) whose deletion eliminates activity (summarized in Fig. 4). Nevertheless, despite this multiplicity of essential or important segments in some S. pombe ARS elements, the inference of Clyne and Kelly (5) may be partially valid. It is possible that, for each ARS element, one essential or important segment may be more important than others. For example, domain α , which can replace but cannot be replaced by domains β and γ in ars3001, may be ultimately more important for ARS activity than domain β or y—even though all three domains are essential in certain tests (Fig. 3B and 4, 5, and 7).

The segments most important for ARS activity are usually abundant in matches to AT-rich asymmetric consensus sequence motifs. Several previous investigators have identified sequence motifs common to S. pombe ARS elements. These sequence motifs share the properties of being highly AT rich and asymmetric (mostly A residues in one strand and T residues in the other strand). We have searched for these consensus motifs in the sequences of the three ARS elements shown in Fig. 4. The positions of the matches we have found to these motifs are shown in Fig. 4 by markers consisting of vertical black lines with horizontal tails. Each marker is labeled to identify the sequence motif matched at that position. The exact locations of matching sequences in ars3001 are shown in Fig. 5. In Fig. 4 and 5, M signifies the 11-bp consensus of Maundrell et al. (25), Z stands for the 12-bp consensus of Zhu et al. (46), C indicates the rightmost 30-bp in segment 1 of Clyne and Kelly (5), and D represents a 20-bp consensus, previously unpublished, that we identified based on the critical sequence in segment 10 of ars3002 (7). In each case in Fig. 4, the direction of the horizontal tail indicates whether the match to the T-rich form of the consensus is found in the upper strand (tail to the right) or lower strand (tail to the left).

It is striking that all clusters consisting of three or more consensus matches are located in DNA segments whose deletion leads to serious loss of ARS activity (Fig. 4). In fact, for the three ARS elements in Fig. 4, each of the 50- to 60-bp segments most important for ARS activity is associated with a cluster of consensus matches, and higher-resolution linker substitution experiments within the important segments reveal that 10- or 20-bp linker substitutions targeting the consensus matches frequently inhibit ARS activity to a greater extent than do linker substitutions in other portions of the important

segments (5, 7) (Fig. 5). Thus, although these and earlier observations indicate that there is no S. pombe ARS consensus sequence with the properties of the ACS in S. cerevisiae (these properties include [i] the existence of a single essential match in each ARS element and [ii] the fact that certain point mutations within the consensus destroy ARS activity), the array of S. pombe ARS consensus motifs tested in Fig. 4 appears to provide a tool with which biologists will be able to predict the important portions of previously uncharacterized S. pombe ARS elements. It seems that a cluster of three or more close matches to these motifs indicates a high probability of importance for ARS function. The increased certainty stems from the addition of ars3001 to the database of relatively well characterized S. pombe ARS elements and from the combined use of all four previously proposed S. pombe ARS consensus motifs. In the future, detailed analyses of additional S. pombe ARS elements will provide even greater certainty.

Although clusters of these four consensus motifs correlate with importance for ARS activity within ARS elements, such clusters cannot be used to predict the locations of ARS elements in *S. pombe* genomic DNA. For example, between residues 1830 and 1925 of the *S. pombe* rDNA sequence flanking ars3001 (GenBank accession no. AF040270), the density of matches to the Z consensus sequence is much higher than that in ars3001, but this stretch of sequence is not part of an ARS element (35). Similarly, despite our relatively advanced understanding of *S. cerevisiae* ARS elements, it is not yet possible to predict their locations in chromosomal DNA.

ars1 contains only a single cluster of three or more consensus motifs (Fig. 4). Presumably the functions carried out by the additional segments containing consensus clusters in ars3001 and ars3002 are handled in ars1 by (probably redundant) segments lacking consensus clusters.

Orientation of consensus matches may be important for ARS activity. Both ars3001 and ars3002 have three separate regions containing clusters of three or more consensus matches. It is interesting that, in both ARS elements, all three clusters have the same orientation (T-rich strand on bottom). The possible importance of this common orientation is suggested by the observation that inversion of domains α and β (and to a lesser extent domain γ) inhibits ARS function (Fig. 7). The probability that such common internal orientation would occur by chance in both ARS elements is relatively high, 1 in 16, so additional *S. pombe* ARS elements with multiple consensus clusters need to be analyzed before the significance of this observation can be fully evaluated.

Role of DNA unwindability in ARS function? Considerable evidence (reviewed in reference 20) suggests that a stretch of easily unwound DNA is frequently an important component of an *S. cerevisiae* ARS element. In addition, an earlier low-resolution study suggested a correlation between ARS elements and easily unwound DNA in *S. pombe* (46). To help determine whether stretches of easily unwound DNA correlate with sequences important for ARS function in *S. pombe*, we used the Thermodyn program (28) to calculate the free energy of unwinding (thin black line) in sliding 100-bp windows across each ARS element in Fig. 4.

In *S. cerevisiae*, a free energy of unwinding of <98 kcal/mol is sufficient for normal ARS function (29). For each of the *S. pombe* ARS elements surveyed in Fig. 4, the free energy of unwinding is <98 kcal/mol across most of the ARS element, probably as a partial consequence of the AT richness of most of the ARS element. By comparison to this *S. cerevisiae* standard, then, DNA unwindability should not be limiting for ARS function in most portions of these *S. pombe* ARS elements, and one would not expect a correlation between local minima in

free energy of unwinding and importance for ARS function. Indeed, no correlation is observed. It is interesting, though, that in *ars3001*, which has the highest average free energy of unwinding of the three studied ARS elements, local minima do fall within two of the three regions most important for ARS activity. Thus, the data in Fig. 4 are partially consistent with an important role for DNA unwindability in ARS function but do not prove or disprove it.

Important regions of *S. pombe* ARS elements. The most important regions of *ars1*, *ars3001*, and *ars3002* all contain clustered consensus sequences (Fig. 4). At least one of these consensus-rich regions, domain α of *ars3001*, can functionally replace two other consensus-rich regions (domains β and γ) (Fig. 7). These observations suggest that all of these important consensus-rich regions may bind a common protein(s) and perform a common function(s).

The fact that, within ars3001 and ars3002, all the consensusrich important regions have the same orientation indicates the possibility of interactions between them. That these interactions are synergistic, not additive, is suggested by the observation that deletion of any one of the consensus-rich important regions in ars3001 (domain α , β , or γ) destroys ARS activity (Fig. 7).

Because initiation of DNA replication in *S. pombe* requires ORC function (11) as in *S. cerevisiae*, it is possible that some of the proteins binding to these important consensus-rich regions are components of *S. pombe* ORC. If so, then some *S. pombe* ARS elements, if not all, may contain more than one ORC binding site. Experiments designed to test this possibility are under way in our laboratory.

ACKNOWLEDGMENTS

We are grateful to J. Aquiles Sanchez for pioneering the study of *S. pombe* rDNA replication in our laboratory; to William Burhans and David Kowalski for comments on the manuscript; and to Debbie Mahoney, Karuna Sharma, and Martin Weinberger for constructive criticism throughout this study.

This research was supported by Public Health Service grant GM49294 from the National Institute of General Medical Sciences and by a grant from the Buffalo Foundation, with additional support from shared resources of the Roswell Park Cancer Center Support, grant P30 CA16056.

REFERENCES

- Balzi, E., A. Di Pietro, A. Goffeau, H. van Heerikhuizen, and J. Klootwijk. 1985. The RNA polymerase I initiation site and the external transcribed spacer of the fission yeast *Schizosaccharomyces pombe* ribosomal RNA genes. Gene 39:165–172.
- Bell, S. P., and B. Stillman. 1992. ATP dependent recognition of eukaryotic origins of DNA replication by a multi-protein complex. Nature 357:128–134.
- Berbee, M. L., and J. W. Taylor. 1993. Dating the evolutionary radiations of the true fungi. Can. J. Bot. 71:1114–1127.
- Broach, J. R., Y.-Y. Li, J. Feldman, M. Jayaram, J. Abraham, K. A. Nasmyth, and J. B. Hicks. 1983. Localization and sequence analysis of yeast origins of DNA replication. Cold Spring Harbor Symp. Qual. Biol. 47:1165–1173.
- Clyne, R. K., and T. J. Kelly. 1995. Genetic analysis of an ARS element from the fission yeast *Schizosaccharomyces pombe*. EMBO J. 14:6348–6357.
- Diffley, J. F. X., and J. H. Cocker. 1992. Protein-DNA interactions at a yeast replication origin. Nature 357:169–172.
- Dubey, D. D., S.-M. Kim, I. T. Todorov, and J. A. Huberman. 1996. Large, complex modular structure of a fission yeast DNA replication origin. Curr. Biol. 6:467–473
- Dubey, D. D., J. Zhu, D. L. Carlson, K. Sharma, and J. A. Huberman. 1994. Three ARS elements contribute to the *ura4* replication origin region in the fission yeast, *Schizosaccharomyces cerevisiae*. EMBO J. 13:3638–3647.
- Fournier, P., C. Gaillardin, L. de Louvencourt, H. Heslot, B. F. Lang, and F. Kaudewitz. 1982. r-DNA plasmid from *Schizosaccharomyces pombe*: cloning and use in yeast transformation. Curr. Genet. 6:31–38.
- Gietz, D., A. St. John, R. A. Woods, and R. H. Schiestl. 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20:1425.
- 11. Grallert, B., and P. Nurse. 1996. The ORC1 homolog orp1 in fission yeast

- plays a key role in regulating onset of S phase. Genes Dev. 10:2644-2654.
- Grimm, C., J. Kohli, J. Murray, and K. Maundrell. 1988. Genetic engineering of Schizosaccharomyces pombe: a system for gene disruption and replacement using the ura4 gene as a selectable marker. Mol. Gen. Genet. 215:81–86.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351–359.
- Holmes, S. G., and M. M. Smith. 1989. Interaction of the H4 autonomously replicating sequence core consensus sequence and its 3'-flanking domain. Mol. Cell. Biol. 9:5464–5472.
- Hsiao, C.-L., and J. Carbon. 1979. High-frequency transformation of yeast by plasmids containing the cloned yeast ARG4 gene. Proc. Natl. Acad. Sci. USA 76:3829–3833.
- Huang, R.-Y., and D. Kowalski. 1996. Multiple DNA elements in ARS305 determine replication origin activity in a yeast chromosome. Nucleic Acids Res. 24:816–823.
- 17. Kim, S.-M., and J. A. Huberman. Unpublished data.
- 18. Kim, S.-M., and J. A. Huberman. Unpublished data.
- 19. Lapeyre, B., and J. Feliu. 1993. GenBank accession no. Z19578.
- Lin, S., and D. Kowalski. 1997. Functional equivalency and diversity of cis-acting elements among yeast replication origins. Mol. Cell. Biol. 17:5473– 5484.
- Liu, Z., A. Zhao, L. Chen, and L. Pape. 1997. Activated levels of rRNA synthesis in fission yeast are driven by an intergenic rDNA region positioned over 2500 nucleotides upstream of the initiation site. Nucleic Acids Res. 25:659–668.
- Maleszka, R., and G. D. Clark-Walker. 1993. Yeasts have a four-fold variation in ribosomal DNA copy number. Yeast 9:53–58.
- Marahrens, Y., and B. Stillman. 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science 255:817–823.
- Matsumoto, K., and Y. Ishimi. 1994. Single-stranded-DNA-binding proteindependent DNA unwinding of the yeast ARSI region. Mol. Cell. Biol. 14: 4624–4632.
- Maundrell, K., A. Hutchison, and S. Shall. 1988. Sequence analysis of ARS elements in fission yeast. EMBO J. 7:2203–2209.
- Maundrell, K., A. P. H. Wright, M. Piper, and S. Shall. 1985. Evaluation of heterologous ARS activity in S. cerevisiae using cloned DNA from S. pombe. Nucleic Acids Res. 13:3711–3722.
- Moreno, S., A. Klar, and P. Nurse. 1991. Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. Methods Enzymol. 194:795–823.
- Natale, D. A., A. E. Schubert, and D. Kowalski. 1992. DNA helical stability accounts for mutational defects in a yeast replication origin. Proc. Natl. Acad. Sci. USA 89:2654–2658.
- Natale, D. A., R. M. Umek, and D. Kowalski. 1993. Ease of DNA unwinding is a conserved property of yeast replication origins. Nucleic Acids Res. 21:555–560.
- Newlon, C. S. 1996. DNA replication in yeast, p. 873–914. In M. L. DePamphilis (ed.), DNA replication in eukaryotic cells. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
- Rao, H., Y. Marahrens, and B. Stillman. 1994. Functional conservation of multiple elements in yeast chromosomal replicators. Mol. Cell. Biol. 14: 7643–7651.
- Rao, H., and B. Stillman. 1995. The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. Proc. Natl. Acad. Sci. USA 92:2224–2228.
- 33. Rowley, A., J. H. Cocker, J. Harwood, and J. F. X. Diffley. 1995. Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. EMBO J. 14:2631–2641.
- Sakaguchi, J., and M. Yamamoto. 1982. Cloned *ura1* locus of *Schizosaccha-romyces pombe* propagates autonomously in this yeast assuming a polymeric form. Proc. Natl. Acad. Sci. USA 79:7819–7823.
- Sanchez, J. A., S.-M. Kim, and J. A. Huberman. 1998. Ribosomal DNA replication in the fission yeast, *Schizosaccharomyces pombe*. Exp. Cell. Res. 238:220–230.
- Schaak, J., J. Mao, and D. Soll. 1982. The 5.8S RNA gene sequence and the ribosomal repeat of *Schizosaccharomyces pombe*. Nucleic Acids Res. 10: 2851–2864
- Sikorski, R. S., and P. Hieter. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics 122:19–27.
- Struhl, K., D. T. Stinchcomb, S. Scherer, and R. W. Davis. 1979. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. Proc. Natl. Acad. Sci. USA 76:1035–1039.
- Theis, J. F., and C. S. Newlon. 1994. Domain B of ARS307 contains two functional elements and contributes to chromosomal replication origin function. Mol. Cell. Biol. 14:7652–7659.
- Toda, T., Y. Nakaseko, O. Niwa, and M. Yanagida. 1984. Mapping of rRNA genes by integration of hybrid plasmids in *Schizosaccharomyces pombe*. Curr. Genet. 8:93–97.
- Umek, R. M., and D. Kowalski. 1988. The ease of DNA unwinding as a determinant of initiation at yeast replication origins. Cell 52:559–567.
- 42. Walker, S. S., S. C. Francesconi, and S. Eisenberg. 1990. A DNA replication

- enhancer in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA 87:4665-
- 43. Wohlgemuth, J. G., G. H. Bulboaca, M. Moghadam, M. S. Caddle, and M. P. Calos. 1994. Physical mapping of origins of replication in the fission yeast Schizosaccharomyces pombe. Mol. Biol. Cell 5:839-849.
- 44. Wright, A. P. H., K. Maundrell, and S. Shall. 1986. Transformation of Schizosaccharomyces pombe by non-homologous, unstable integration of
- plasmids in the genome. Curr. Genet. **10:**503–508. 45. **Zhao, Y., and H. B. Lieberman.** 1995. *Schizosaccharomyces pombe*: a model for molecular studies of eukaryotic genes. DNA Cell Biol. 14:359-371.
- 46. Zhu, J., D. L. Carlson, D. D. Dubey, K. Sharma, and J. A. Huberman. 1994. Comparison of the two major ARS elements of the *ura4* replication origin region with other ARS elements in the fission yeast, *Schizosaccharomyces* pombe. Chromosoma 103:414-422.